CSDQ: A USER-CENTERED APPROACH TO IMPROVING THE QUALITY OF CUSTOMER SUPPORT DATA

(Practice-Oriented, Methodologies)

Susan L. Keenan
Hewlett-Packard Company, Littleton, MA, USA
susan.keenan@hp.com

Terri Simmons
Hewlett-Packard Company, Houston, TX, USA
terri.simmons@hp.com

Abstract: Improving data quality in a customer support/business environment presents unique challenges to information quality (IQ) teams. In these environments, data may be collected and used across a worldwide network of people, systems, and processes; and, the success of a quality improvement program depends largely on participation and buy-in across organizational and regional boundaries. To incorporate the human factor in data quality improvement programs, existing methodologies must be augmented to fully address the needs and constraints of multiple, diverse groups of data consumers. This paper presents the Customer Support Data Quality (CSDQ) method, a user-centered approach to improving the quality of customer support data. CSDQ extends the Total Data Quality Management methodology by integrating user-centered methods into the continual data quality improvement framework. A case study is presented that describes how CSDQ was applied in a customer support environment at Hewlett-Packard Company.

Key Words: Data Quality, Information Quality, Qualitative Data Analysis, Customer Support

INTRODUCTION AND BACKGROUND

Data gathered in customer support environments are an integral component of corporate data collection programs – these data contribute significantly to corporate-wide strategies on warranty, support contract negotiations, and product/service offerings. The quality of this information (we use the terms "data" and "information" interchangeably throughout this paper) not only affects customer satisfaction/loyalty, operational efficiency, and employee satisfaction, but impacts decision-making, strategy, and the level of trust among organizations [4][7]. Current recommendations suggest that, to ensure high-quality information, companies establish a continual assessment and improvement program that includes specific activities for defining information products/metrics, assessing the level of quality, and analyzing root causes of poor quality data. While these methodologies are well grounded in both research and practice, they focus on information products, not on information users. For global companies that collect and use data across a worldwide network of people, systems, and processes, new methods are needed that integrate the human factor into a data quality program. Specifically, methods are needed that encompass factors related to users, their tasks, and their cultural/work environments – methods that increase the level of participation across regional and organizational boundaries, and support communication of improvement strategies to stakeholders and sponsors.

The flow of information in a typical customer support/business environment is essentially a loop (see Figure 1). The flow begins when customers call in and talk to agents about issues and questions. Agents capture information about those calls in one or more databases. Reporters analyze the data and report their findings to engineers who develop recommendations about products and services. The recommendations are passed to the business units, where decisions are made about fixes and product/service offerings.

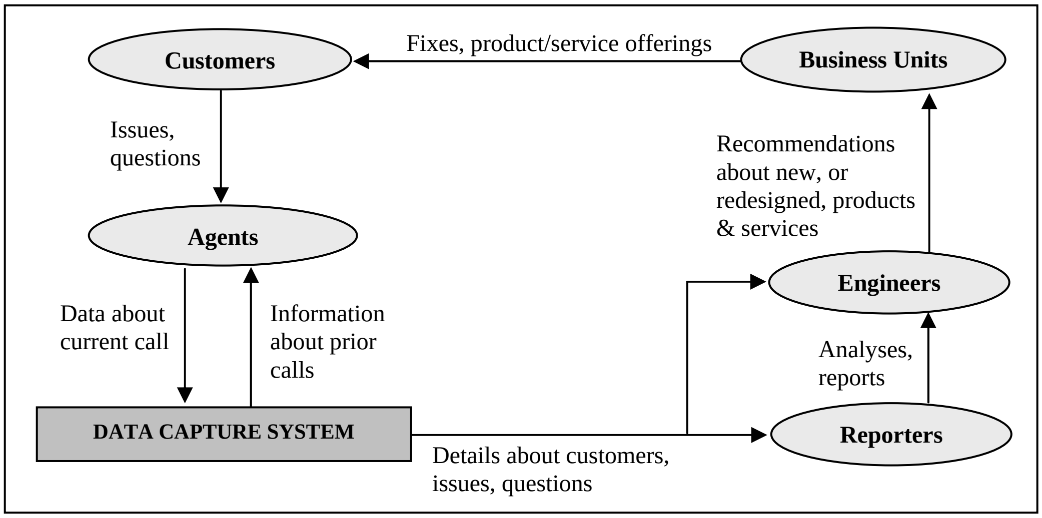

Figure 1. Information flow in a customer support/business environment.

All five groups of individuals, shown in the ovals in Figure 1, are impacted by the quality of the information stored in the data capture system. Missing or inaccurate data have an immediate affect on the support experience – when agents repeatedly ask customers for the same information, call length is increased, and customers become dissatisfied. Agent satisfaction and performance also suffer, and the support center may not meet its satisfaction and performance goals. Poor quality data also limit the effectiveness of text mining applications and automated analyses. Reporters must read call details and manually bucket unreliable data as they look for emerging issues and trends. Engineers also read call details to explain or verify information in the reports. The adage "garbage in, garbage out" applies: engineers need high quality information on which product/service recommendations can be based with assurance. Poor quality information can contribute to business decisions that have a negative impact on customer satisfaction, loyalty, warranty costs, operational efficiency, and sales.

The growing costs associated with poor quality data have fueled much research in quality assessment and improvement. Total Data Quality Management (TDQM) outlines four life cycle phases in a continuous improvement program (define, measure, analyze, and improve) and emphasizes the need to establish an information quality (IQ) team (with a senior executive as the champion), teach IQ assessment and management, and institutionalize continuous information improvement [9]. Wang and Strong present a hierarchical framework consisting of data quality dimensions that are important to data consumers [8]. The AIMQ methodology for assessing and benchmarking quality uses the product and service performance model for information quality (PSP/IQ), the IQA instrument, and IQ Gap Analysis techniques [2]. Willshire and Meyen describe the Data Quality Engineering Framework (DQEF), which can be tailored by organizations to meet specific needs related to measuring data quality and developing improvement plans [10]. The Small Business Quality Framework (SBQF) extends and modifies TDQM to reduce resources and lower costs associated with assessing and improving data quality in small business environments [3]. Use-based data quality programs focus on improving the use of critical data and incorporate use-based data quality audits, redesign, training, and continuous measurement [5]. And, Pipino et al. outline principles that can be used to develop usable data quality metrics [6].

The application of existing data quality methodologies in a customer support/business environment is challenging for several reasons. First, there are four dissimilar groups of data consumers (agents, reporters, engineers, and business units) with diverse needs and goals regarding data collection and usage. For example, agents focus on customer satisfaction and issue resolution – not on data collection. Yet agents often need to review information captured about prior calls; reporters need data that accurately reflect the issues; engineers need "actionable" information (i.e., information that can be used to reproduce the issue, identify the root cause, and develop corrective action plans); and, business units need data that can guide fixes, product/service offerings, and operational decisions. Second, agents often capture data about the issue, troubleshooting, and resolution by typing prose in free-form text fields. This type of data is difficult to measure or validate. In addition, the level of detail an agent provides often depends on his/her experience level, workload, and preference for recording issue symptoms or causes. Third, for global companies, regional variations in business processes and data collection procedures can introduce significant differences in the type and quantity of information in a given field.

We believe that integrating user-centered methods in the data quality assessment and improvement process will enable IQ teams to meet the challenges outlined above. User-centered investigation methods are especially suitable for characterizing multiple, diverse user groups, identifying the needs of those groups, and prioritizing the needs in the context of organizational, environmental, and ethnographic concerns. User-centered design and evaluation methods are, likewise, valuable tools for improving screens that disrupt the agent workflow and have a negative impact on data quality. Our experiences indicate that by focusing on data users as well as the quality of the output, IQ teams can formulate comprehensive improvement strategies that work within the constraints of existing business processes.

This paper presents the Customer Support Data Quality (CSDQ) method, a user-centered approach to improving the quality of customer support data. Section 2 outlines the CSDQ method. Section 3 presents a case study conducted at Hewlett-Packard Company in which CSDQ was applied in a customer support environment. Section 4 concludes with lessons learned during the case study.

THE CUSTOMER SUPPORT DATA QUALITY (CSDQ) METHOD

In the Customer Support Data Quality (CSDQ) method, user-centered methods are combined with proven data quality methods to ensure that data quality assessment and improvement are considered within the context of key factors related to people, systems, and processes. CSDQ is an iterative process, designed to work within the continual data quality improvement framework outlined in the Total Data Quality Management methodology [9]. See Figure 2.

Although CSDQ follows the four Total Data Quality Management phases (define, measure, analyze, and improve), it extends the quality improvement framework in several ways. First, the quality team focuses on four areas:

- Information – for which assessments of quality may vary by user group,
- People – who may be inadequately trained, lack understanding, or have conflicting needs,
- Systems – which may be difficult to use, awkward, or unreliable, and
- Processes – which may vary across organizations and contain bottlenecks, communication gaps, and ineffective procedures.

Second, user-centered methods are integrated into each phase. These methods can be used to support data gathering and definition, broaden the focus of measurement and analysis, and integrate a user-centered perspective in the formulation of improvement strategies. Examples of user-centered methods that can be used during each phase are shown in the gray regions in Figure 2. Third, a path is added from the improvement phase to the measurement phase. The new path emphasizes the benefit of repeating the measurement, analysis, and improvement cycle prior to implementing final solutions. The path also provides a means for assessing acceptance (buy-in) of proposed improvement strategies – which is critical when solutions must be implemented across regions, business units, and/or organizations. Last, CSDQ introduces four quick-and-dirty analysis techniques (the tallies table, the causes table, the causes diagram, and the dependency timeline) which incorporate ideas from defect causal analysis [1], a method for improving software quality by analyzing the causes of systematic errors. These techniques are shown in the rounded brackets in the gray areas at the bottom of Figure 2. CSDQ phases, methods, and outcomes are discussed below.

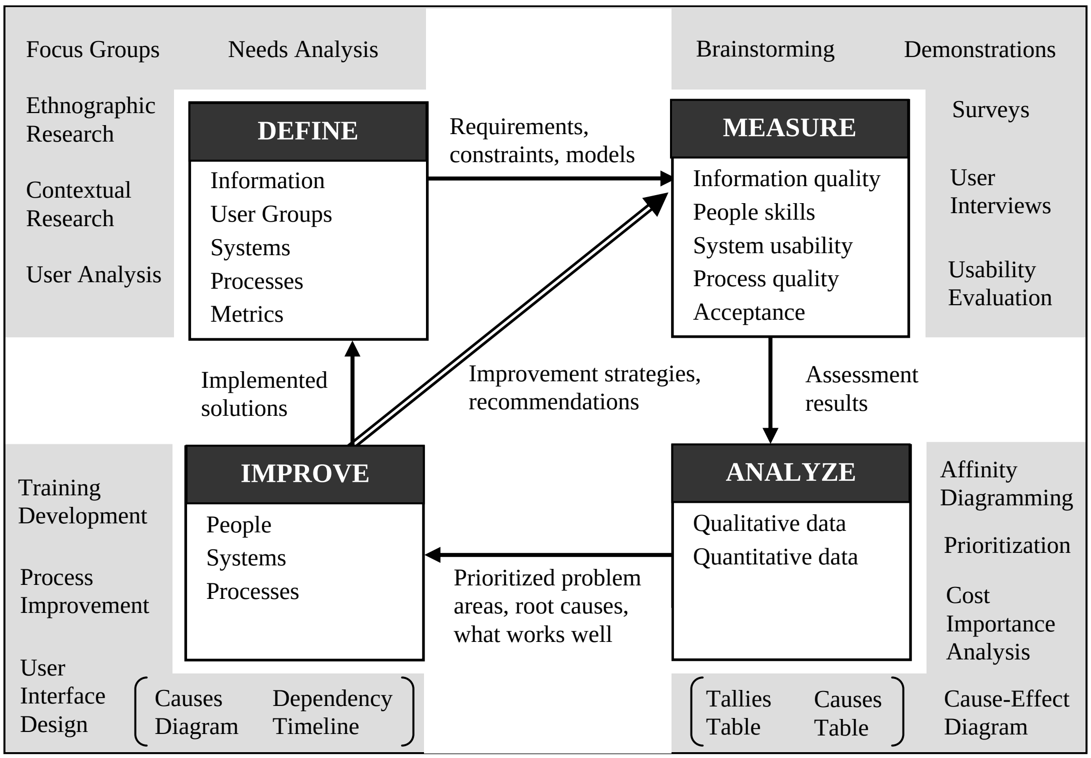

Figure 2. The CSDQ method.

Define

During the definition phase, user-centered research methods are conducted to define information products and IQ dimensions; profile user groups, stakeholders, and organizations; explore needs related to existing systems and processes; and identify relevant metrics. Five user-centered definition methods are highlighted in the upper left gray region in Figure 2 (needs and user analysis, ethnographic and contextual research, and focus groups).

The outcomes of the definition phase include requirements, constraints, and an understanding of existing systems (e.g., the data capture tool, databases, and data analysis/reporting tools) as well as business practices (e.g., procedures for training, data capture, analysis, and reporting). It is helpful to develop a set of models that define the project space and document factors that may impact the formulation of improvement strategies. These models can include data flow diagrams, user models that capture demographic and task information about each user group, and models that document communication flow.

Measure

During the measurement phase, user-centered methods are used in conjunction with proven data quality methods to measure the level of quality and uncover problem areas for each user group. Five user-centered assessment methods are highlighted in the gray region in the upper right corner of Figure 2 (brainstorming, user interviews, surveys, usability evaluation, and demonstrations).

The choice of methods depends on the goals of the assessment, the availability of participants, and the point in the project at which the measurement (assessment) occurs. For example, consider a situation where the IQ team performs the measurement, analysis, and improvement cycle twice. During the first iteration, the team may use surveys and interviews to gather data and measure problem areas. During the second iteration, the team may administer follow-up surveys to assess improvement and perform demonstrations to assess the level of buy-in.

Assessment outcomes yield both qualitative and quantitative data. These data are typically related to information quality, people, systems, and/or processes, but can also include metrics that surface during assessment as well as indicators of the level of acceptance/buy-in of proposed solutions.

Analyze

The primary goals of the analysis phase are to assess the impact of each problem area and identify root causes. CSDQ draws on a range of analysis methods that can be applied to qualitative data. Three user-centered methods (affinity diagramming, cost/importance analysis, and prioritization) can be used to group issues, assign costs and importance ratings, and prioritize issues. The cause-effect diagram (also known as the Ishikawa or fishbone diagram) can be used to sort causes that contribute to a specific problem area. We also include two quick-and-dirty techniques (the tallies table and the causes table) that we developed during the case study project to help us summarize our analyses and communicate our findings to a larger group.

In this section, we focus on the analysis of qualitative, interview data as the literature contains many techniques for analyzing quantitative data. We recommend a two-step approach to analyzing qualitative data:

1. Sort and tally qualitative data.
2. Identify, and rate the importance of, primary data quality attributes. Identify the causes of each primary attribute and assess the significance of each cause.

Each step is described below.

<u>Step 1. Sort and tally qualitative data.</u>
To sort and tally qualitative data, first sort the data using four buckets: data quality, people, systems/tools, and process. Then sort the data within each bucket into groups called attributes. Tally and record the number of issues in each attribute for each user group.

We use the tallies table, shown in Table 1, to capture the results of this step. The leftmost “Bucket” column contains the four buckets. The “Attribute” column contains the attributes that emerge when the data within each bucket are sorted. The tallies for each user group are recorded in the “User Groups” columns (the actual counts are represented by “X” due to the proprietary nature of the case study data).

BUCKET	ATTRIBUTE	USER GROUPS		
		Agents	Reporters	Engineers
Data Quality	Believability	X	X	X
	Accuracy	X	X	X
	Reputation	X	X	X
	Relevancy			X
	Value-added			X
	Completeness	X	X	X
	Appropriate amount of data	X	X	X
	Interpretability		X	X
	Ease of understanding		X	X
	Consistent representation	X	X	X
	Accessibility	X		
People	Agent training	X		
	Agent turnover rate		X	X
System / Tool	Usability	X		
	Reliability	X		
	Functionality	X	X	X
	Performance	X		
	Language	X	X	X
Process	Data capture procedures	X		
	Manual workarounds	X	X	X
	Call center business processes	X		
	Call center performance metrics	X		

Table 1. Tallies table for the case study.

We recommend that the data in the data quality bucket be sorted using the 15 quality dimensions outlined in the conceptual framework of data quality [8]. These dimensions were developed empirically and provide coverage for aspects of data quality that are important to data consumers [2][8]. Using these dimensions also facilitates the comparison of qualitative interview data with quantitative data collected using the IQA instrument. When the case study data were sorted, issues in the data quality bucket were sorted into 11 of the 15 data quality dimensions (these 11 dimensions are the attributes in the data quality bucket in Table 1). When the issues in the people bucket were sorted, two attributes emerged, five emerged from the system/tool bucket, and four emerged from the process bucket. Specific attributes will be discussed in detail in section 3.

<u>Step 2. Identify, and rate the importance of, primary data quality attributes. Identify the causes of each primary attribute and assess the significance of each cause.</u>

Label each data quality attribute as either a primary or a secondary attribute. To identify primary attributes, think about whether a given attribute depends on other attributes, and how soon a change in the level of quality might be observed. For example, consider accuracy and believability. If the level of accuracy does not depend on other attributes, it is a primary attribute. If believability depends on accuracy, believability is a secondary attribute. Improvements in accuracy would likely be observed soon after a change has been implemented, whereas improvements in believability may not be observed as quickly.

Next, rate the importance of each primary attribute. The rating can be based on variables such as severity, cost, and impact. To identify the causes of each primary attribute, think about whether the issues classified in the attribute are due to people, tools, and/or processes. (One source of causes are the issues classified in the people, system/tool, and process buckets in Table 1.) Record the causes of each primary attribute; assess the significance of each cause, and highlight the most significant cause.

We use the causes table, shown in Table 2, to summarize the causal analysis. Each of the 11 data quality attributes identified in step 1 has been labeled as either a primary or a secondary attribute. Five of the 11 attributes are primary attributes; six are secondary attributes. The importance rating for each primary attribute is in the "Attribute Importance" column. A "✓" is used to indicate a cause. A "✓✓" indicates the most significant cause for an attribute, if one exists. For example, consider the accuracy and completeness attributes in Table 2. The most significant cause of accuracy issues is tools and the most significant cause of completeness issues is process.

DATA QUALITY ATTRIBUTE	PRIMARY OR SECONDARY	ATTRIBUTE IMPORTANCE	CAUSED BY		
			People	Tools	Process
Believability	Secondary				
Accuracy	Primary	High	✓	✓✓	
Reputation	Secondary				
Relevancy	Secondary				
Value-added	Secondary				
Completeness	Primary	High	✓	✓	✓✓
Appropriate amount of data	Primary	Medium	✓✓	✓	✓
Interpretability	Secondary				
Ease of understanding	Secondary				
Consistent representation	Primary	Low		✓	
Accessibility	Primary	High		✓	

Table 2. Causes table for the case study.

Improve

During the improvement phase, the IQ team formulates improvement strategies to address the causes of the primary data quality attributes. Improvement activities vary depending on the causes that need to be addressed, e.g., addressing a people cause might involve developing new training materials, addressing a process cause may involve implementing a new process, and addressing a system cause could involve redesigning the user interface. The outcomes of the improvement phase include improvement strategies and implemented solutions.

Two additional diagrams can help focus and set expectations about solution strategies: the causes diagram and the dependency timeline. The causes diagram is a Venn diagram that illustrates the relationship among the primary attributes and causes in the causes table. The causes diagram for the case study is shown in Figure 3. The attributes assessed at high importance are underlined. Most significant causes are in the parentheses, e.g., Accuracy (Tools) indicates that tools were the most significant cause of accuracy issues. Although there is no new information in the causes diagram, we found that it clearly indicates the areas on which the improvement efforts should focus and thus could be used to communicate the rationale behind proposed improvement strategies to sponsors and stakeholders.

A dependency timeline can be used to guide initial reassessment and set expectations regarding when improvements in specific attributes are likely to be observed. In some cases, the dependency timeline may help further clarify the relationship among attributes. Consider the dependency timeline for the case study, shown in Figure 4. After additional dependencies among the secondary attributes were identified, all attributes were mapped to an improvement timeline. The timeline will be discussed in more detail in section 3.

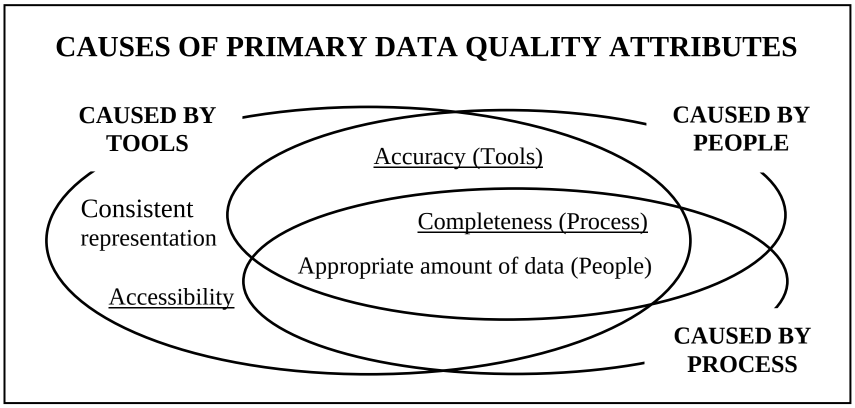

Figure 3. Causes diagram for the case study.

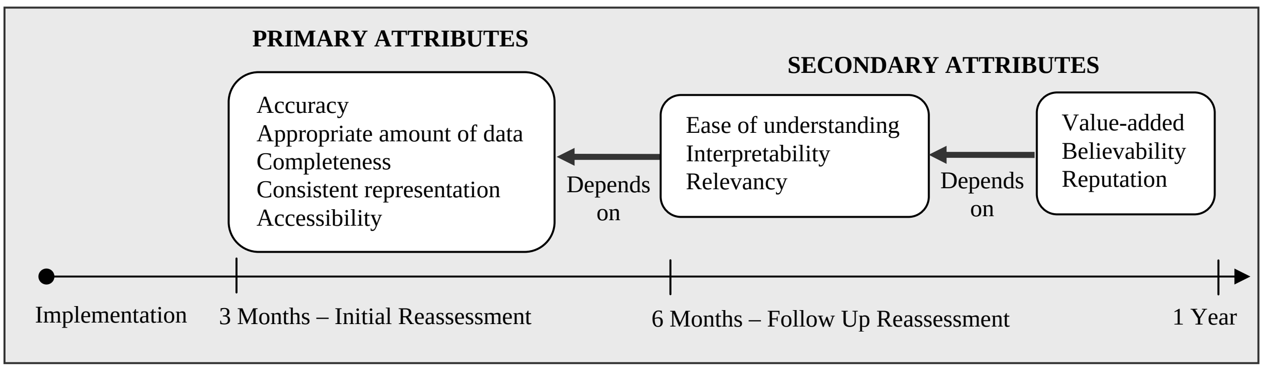

Figure 4. Dependency timeline for the case study.

THE CASE STUDY

The project described in this section was initiated on behalf of Hewlett-Packard's consumer imaging and printing product groups. The primary goal of the project was to improve the accuracy of the information captured by call agents about customer issues. The project spanned a seven-month period.

Resources

The IQ team was comprised of two human factors consultants. The team devoted approximately 50% of their time to the project during the first few months, significantly less time during the latter months. Other resources included periodic, nominal access to members of each user group.

Applying CSDQ to the Project

The team performed the definition phase once and then iterated through the measurement, analysis, and improvement cycle twice. The first iteration focused on investigating needs related to information, people, systems/tools, and processes, and on formulating improvement strategies to address those needs. The second iteration focused on assessing the proposed solutions and the level of acceptance with which those solutions were received.

Define Phase

During the definition phase, the IQ team gathered background information on data/information, people, systems/tools, and process:

- <u>Data/Information</u>. The IQ team reviewed data collected about individual customer issues as well as data summaries in weekly/monthly reports. The data captured about each customer issue included a prose description typed in a large text field and an issue classification obtained by navigating a large, hierarchical classification scheme. Outcomes of this review included a list of information products and the one quality metric discussed by stakeholders – accuracy. In this context, accuracy was defined as the degree to which the agent's classification of a customer problem matches the actual problem in the prose description.

- <u>People</u>. The IQ team identified stakeholders and three distinct groups of data users (agents, reporters, and engineers). The team performed user analysis to understand the characteristics of each group and needs analysis to develop a preliminary understanding of the most important needs of each group.

  People related outcomes underscored the diverse nature of the three user groups – there were significant differences in background and experience, data-related tasks, and performance and business metrics. *Agents* are trained on the tools they will use and on the product(s) to which they are assigned. An agent's primary task is to resolve customer issues as quickly as possible. Data-related tasks include capturing information during a call and reviewing information captured on previous calls. Typical agent metrics are average call handling time, customer satisfaction measures, and first-call resolution rate. *Reporters* are often referred to as the intermediaries between the call centers and HP product divisions. Many reporters are former agents. Their primary tasks are to analyze call data and produce reports that identify emerging issues and trends. Reporters may spend several hours per week validating the accuracy of call data. *Engineers* work within the product R&D groups and use the information contained in the reports to drive product quality improvements.

- <u>Systems/Tools</u>. The IQ team researched available documentation on the systems/tools used by each user group. Of particular interest were the tools used by agents to capture data, look-up entitlement, warranties, and part numbers, and search databases containing troubleshooting and solution documents.

- <u>Process</u>. The team examined communication processes, business processes, and procedures related to data capture, analysis, reporting, and training. Process related outcomes included the data flow diagram shown in Figure 1, notes documenting regional variations in business practices, and a model of the agent workflow. The agent workflow model outlined the four phases of each customer call (collect background information, identify the reason for the call, troubleshoot and resolve the issue, and wrap up) and documented the type of information captured during each phase.

The information gathered during the definition phase highlighted why previous efforts to improve accuracy, which had focused primarily on the needs of engineers, had been largely unsuccessful. The entire agent experience impacts an agent's ability to capture high-quality information. Previous attempts to improve accuracy by modifying the hierarchical classification scheme overlooked other important factors related to agent needs, tools, and processes. The outcomes of this phase emphasized the need to refocus the project on agents and data capture, thus driving activity selection in the remaining project phases.

Iteration #1 (Measure, Analyze, Improve)

Measure Phase

During the assessment phase, global user interviews were conducted with over 70 members of the three data user groups (agents, reporters, and engineers). A script was developed for the interviews that focused on:

- Demographics (e.g., job role, responsibilities, product lines supported, and region),
- User tasks (e.g., how each user group touches information, how frequently each task is performed, and the time it takes to perform each task),
- Needs of each user group,
- The type of information that needs to be captured about customer issues (e.g., issue description, troubleshooting steps and issue resolution, and issue classification),
- Tools,
- Existing business processes,
- Information transfer among user groups,
- Analysis and reporting,
- Metrics, and
- The biggest frustration, the one change to make, and the benefit of that one change.

The interviews were conducted remotely, with 3-15 participants in each interview session. The team also interviewed a few call center managers who provided information about business processes at the call centers, characterized the relationship between the call center and HP, and described the metrics against which agents and call centers are measured.

Analyze Phase

During the analysis phase, the IQ team followed the two-step process outlined in section 2 to analyze the qualitative interview data. In step 1, the issues in each bucket were sorted and tallied for each user group. The issues in the data quality bucket were categorized using the 15 dimensions suggested by Wang and Strong [8]. These attributes provided complete coverage for the data quality issues, i.e., there were no "leftover" issues at the end of the sorting process. Four of the 15 attributes (objectivity, timeliness, concise representation, and access security) were not used and thus were not included in Table 1. When the issues in the people bucket were sorted, two attributes emerged: inconsistent training given to agents and the high turnover rate among agents. When the issues in the system/tool bucket were sorted, five groups emerged: poor system usability, inadequate reliability, missing features, poor system performance (e.g., slow response), and language issues (e.g., unclear terminology and lack of translation). Note that a clear distinction was made between system/tool issues and the issues categorized in the accessibility attribute in the data quality bucket. When the issues in the process bucket were sorted, four attributes emerged: regional variation in data capture procedures and call center business processes, the need for manual workarounds, and the impact of agent performance metrics on data capture. Some examples of comments and the buckets in which they were placed are shown in Table 3.

In step 2, the IQ team first labeled each data quality attribute as a primary or secondary attribute. Five primary attributes were identified: accuracy, completeness, appropriate amount of data, consistent representation, and accessibility. Next the team assigned an importance rating to each primary attribute. Accuracy was clearly of high importance – it was the one aspect of data quality that sponsors mentioned at the onset of the project. Completeness and accessibility were also assigned high ratings – incomplete data impacted every interview participant and accessibility issues often prevented agents from capturing data until the end of their shifts. Appropriate amount of data and consistent representation were assigned ratings of medium and low, respectively. See Table 2.

USER GROUP	BUCKET	ATTRIBUTE	COMMENT
Engineer	Data Quality	Accuracy	"The (call) details don't match the classification"
Engineer	Data Quality	Completeness	"Not all of the … steps (taken by the agent) are documented"
Agent	People	Agent training	"…confusion about what the hierarchy is for and how to use it"
Agent	System/Tool	Usability	"The tree structure has too many levels"
Agent	System/Tool	Performance	"Speed of the tool – it takes too long to refresh"
Reporter	System/Tool	Language	"Agents cannot provide (write) details in English"
Agent	Process	Manual workarounds	"I capture information in Notepad"

Table 3. Sample comments from the interviews.

Next, using the people, tools, and process buckets, the team identified the causes of each primary attribute and assessed the significance of each cause. Accuracy issues were caused by people (many agents did not understand the benefit of accurate classification) and tools (the classification hierarchy was difficult to navigate and classification categories were not translated into local languages). Issues related to completeness and appropriate amount of data were caused by people, tools, and processes. Completeness issues were primarily due to the differences in call center processes and issues related to appropriate amount of data were caused primarily by inconsistent training. Issues related to consistent representation and accessibility were both caused by tools. The team used the number of comments made by members of each user group and the emphasis with which those comments were made to determine the most significant causes. See Table 2.

<u>Improve Phase</u>

During the improvement phase, the team created a causes diagram (recall Figure 3). The diagram emphasized that redesigning the data capture tool was critical to the success of the project; however, it was clear that a complete solution strategy must also encompass improvements related to people and process.

Requirements and constraints for the new data capture model were based on the issues classified in the people, system/tool, and process buckets. The most important requirements were:

- Support the agent workflow by aligning data capture with the flow of a call.
- Facilitate data capture for agents for whom English is a second language.
- Minimize/eliminate system performance issues related to user interface design.
- Reduce the need for manual workarounds.
- Ensure that agents can continue to meet performance metrics.
- Increase the accuracy with which agents classify customer issues.
- Ensure adaptability across product lines.

The new model was prototyped using HTML and JavaScript. The model enabled agents to select checkboxes rather than type all call details in one large, unformatted text field, and to create a classification from the checkboxes that had been selected rather than navigate a hierarchy to classify each issue. A new call-details summary was automatically formatted by the prototype to facilitate scanning and data mining. In addition, three new information products identified during the definition phase were introduced into the design.

Iteration #2 (Measure, Analyze, Improve)

Measure Phase

During the second iteration of the measurement phase, multiple evaluation techniques were performed on the prototype. Demos were conducted remotely across the regions for stakeholders, sponsors, and members of each user group. The demos were designed to gather feedback and assess the level of acceptance of the new data capture model. Brainstorming, usability testing, and satisfaction questionnaires were conducted on site at three call centers with agents and supervisors. Brainstorming sessions were designed to elicit feedback from agents and supervisors about the new data capture model and the degree to which the model would support agent tasks, metrics, and existing call center business practices. The usability tests were designed to assess the level of usability achieved in the prototype, identify any remaining usability issues, and measure agent satisfaction. User satisfaction questionnaires were used to assess the level of satisfaction with the prototype user interface and the new data capture model.

The team administered two different satisfaction questionnaires. The first questionnaire was administered at the conclusion of the brainstorming sessions, and contained standard user satisfaction questions that employed a 10-point scale to assess satisfaction with learnability, usability, and memorability. The second questionnaire, which was administered during the usability testing sessions, contained questions about specific aspects of the prototype user interface and the utility of the new data capture model. Both types of questionnaires were completed manually.

During each testing session, the participating agent would first work through a series of simulated customer calls (which were based on transcripts of actual calls), and then complete the second satisfaction questionnaire discussed above. To simulate the calls, an IQ team member played the role of the customer and the participating agent attempted to resolve the customer issue, using the prototype to capture information about the customer, the issue, troubleshooting steps, resolution, and issue classification. The prototype automatically saved the data entered by the agents and stored a time stamp when a page loaded.

Both qualitative data and quantitative data were collected during the assessment activities. Qualitative data were gathered during the demos, brainstorming sessions, and usability tests. These data included positive and negative feedback, concerns, usability issues, and the data entered by participating agents during the usability testing sessions. Quantitative usability data included performance measures (time on task and time on page) and data gathered from the questionnaires.

Analyze Phase

During the second iteration of the analysis phase, the IQ team explored several questions:

- Is issue classification more accurate with the new model?
- Does the prototype achieve a high degree of usability and user satisfaction?
- Do the new call summaries improve analysis and reporting?
- What is the level of acceptance/buy-in of the new data capture model?

The qualitative and quantitative evaluation data provided strong evidence that the model is an effective tool for capturing high-quality data. As to the first question, the accuracy with which the agents classified issues during the testing sessions reflects a 26% improvement over the baseline measurement for the existing classification hierarchy (see Table 4). We anticipate that accuracy will improve with regular usage and training (agents received only 5-10 minutes of training on the prototype at the start of each testing session). Secondly, a high level of usability and user satisfaction was achieved in the prototype. Usability issues observed during testing were minor and agents were able to complete each simulated call. Satisfaction ratings for learnability, usability, and memorability (obtained from administering the standard user satisfaction questionnaire at the conclusion of each brainstorming session) were 8.5., 8.0, and 8.1 on the 10-point scale, respectively. Responses to the second questionnaire (administered to agents at the end of each testing session) were also positive. Thirdly, the new call summary, which contains all data entered for an individual call, is formatted to facilitate scanning and data mining. Lastly, the analyses indicated a strong level of buy-in. Stakeholders, sponsors, and members of each user group were enthusiastic about the novel aspects of the new model, were satisfied with how easy the screens in the prototype were to use, and believed that the model would be readily adopted across regions.

CLASSIFICATION MODEL	ACCURACY
Baseline accuracy with existing hierarchical classification scheme	70%
Accuracy for new, prototyped data capture model	96%

Table 4. Demonstrated improvement in level of accuracy.

Since the team was not given the bandwidth to collect “before” usability data, a before/after comparison of quantitative usability metrics was not possible. Instead the team compared usability measures gathered during the testing sessions (average time on task and time on page) against existing agent performance metrics such as average call handling time. In retrospect, the comparison against the performance metrics seemed to be the more important comparison for stakeholders, sponsors, and call center supervisors. In addition, comments gathered during the demos, brainstorming sessions, and usability testing were compared against the comments made during the initial interviews. Comments about the new data capture model were very positive and indicated that many of the issues raised during the interviews had been addressed in the prototype.

Improve Phase

During the second iteration of the improvement phase, the team addressed the remaining usability issues, documented the total solution, created a dependency timeline (recall Figure 4), and outlined the plan going forward. After making minor modifications to the prototype, the team drafted a user interface specification and created sample reporting templates that highlighted new reporting capabilities. Recommendations for training were developed that described new features, new information products, and data collection procedures. The recommendations also emphasized the importance of supplying accurate data to HP and outlined the benefits of the new model to the agent (e.g., the new screens support the flow of a call, reduce the need for manual workarounds, improve performance by minimizing the number of hits to the server, and reduce reliance on typing information in English).

The plan going forward is focused on the need to reassess data quality following implementation and establish a continual assessment process. As shown in Figure 4, initial reassessment will begin three months following implementation. Members of each user group will be asked to complete a questionnaire. The questionnaire for each group will contain questions from the IQA instrument that are about the primary categories mentioned by that group during the user interviews. For example, the questionnaire given to reporters and engineers will contain questions about four primary attributes (accuracy, appropriate of data, completeness, and consistent representation). Similarly, the questionnaire given to agents will contain questions on all five primary attributes. Follow-up reassessment will be conducted six months to one year following implementation. The team is considering two follow up activities: (A) administer a survey containing questions about the primary attributes, ease of understanding, interpretability, and relevancy at the six month mark; and, (B) administer a survey containing questions about all 11 attributes closer to the one year mark. Depending on resource constraints (A) may be omitted. As in the initial assessment, the survey administered to each user group will contain questions about the attributes mentioned by that group. The plan to establish a continual assessment process is under consideration.

OBSERVATIONS AND CONCLUSIONS

CSDQ's blend of methods from two different domains, human factors and data quality, has proven to be an effective approach for assessing and improving data quality in the context of a world-wide customer support/business environment. In addition to incorporating user-centered methods within the Total Data Quality Management methodology, CSDQ introduces four quick-and-dirty techniques. The tallies table and causes table enable IQ teams to identify the most important data quality attributes for each user group and summarize the causes for each attribute. The causes diagram can help IQ teams focus the scope of improvement strategies. And, the dependency timeline can be used to guide initial reassessment activities and set expectations regarding when improvements are likely to be observed.

When CSDQ was applied to the case study project, improvements were observed in several areas. As mentioned in section 3, a 26% improvement (over the baseline measure) was noted in the accuracy of the issues classified during user testing. The agent experience was improved by easing the burden of data capture. The new model supports the agent workflow by enabling agents to begin classifying as they gather background information and troubleshoot issues – instead of classifying an issue at the end of a call. In addition, the relatively flat classification structure (i.e., lists of checkboxes) is easier to navigate and serves as a prompt for agents as they troubleshoot. Additional benefits are outlined in the lessons learned below:

- <u>Do some interviews</u>. The interviews were more critical to the success of the project than we had originally anticipated. The interviews gave members of each user group a voice in the assessment process, provided visibility for regional representatives, and generated a high level of enthusiasm for the project as well as the CSDQ process. The interviews added richness to the details – word choice and tone-of-voice were excellent indicators of how strongly a speaker felt about a given topic. Responses to questions such as "the one thing to change" and "the benefit of that change" helped the team identify top priorities. In addition, the interviews enabled the team to expand the definition of data quality from "just accuracy" to include other aspects of data quality that were also important to the participants.

- <u>Conduct some activities onsite</u>. In the case study, the team went onsite at three call centers to conduct evaluation activities. The team learned much by observing the physical space, the agent workflow, and the professionalism of the agents. For example, cubicles open onto other cubicles. This setup enables agents to put customers on hold so that they can discuss issues easily with other agents. Agents are not always assigned cubicles and instead select vacant cubicles at the start of their shifts. Agents without assigned cubicles carry specific artifacts to and from work. Artifacts that remain in the cubicles (e.g., top 10 issue list for the month) are placed there by call center managers. Call centers also have large display areas for employee awards, e.g., top call agent for the week. Agents take pride in receiving the awards and are very enthusiastic about representing HP. This type of contextual knowledge is invaluable, and would have been very helpful to the team during the definition phase. Even though the team did not visit the call centers until evaluation, they were able to use this knowledge as they made improvements to the prototype and the data capture model.

- Be flexible about method selection. Although specific activities may be outlined and agreed to at the onset of the project, it is important to be flexible about which specific, user-centered methods will be performed during each phase. During the case study project, the IQ team reassessed the methods that were planned at the start of each project phase. Occasionally the team added/subtracted a method or adapted the way in which a method would be performed.

- Set reasonable improvement goals. In addition to identifying the attributes that must be improved, it is also important to identify attributes and causes that cannot be changed. For example, in the case study, the IQ team was not going to be involved in final implementation, and thus had no control over the technology used to implement the new screens. As a result, they could only make recommendations regarding accessibility and performance. The team also could not effect changes in agent turnover, performance metrics, and some business processes.

- Obtain buy-in. Two parts of the process were critical to obtaining global support for the project: the user interviews and the multiple evaluation methods (demos, brainstorming, and usability testing). On several occasions, participants in these activities said that they were excited about seeing their comments reflected in the new designs.

REFERENCES

[1] Card, D. N. Learning From Our Mistakes With Defect Causal Analysis, *IEEE Software.* 15(1), 1998, pp: 56-63.

[2] Lee, Y. W., D. M. Strong, B. K. Kahn, and R. Y. Wang, AIMQ: a methodology for information quality assessment, *Information & Management.* 40 2002, pp: 133-146.

[3] Leonowich-Graham, P. and M. J. Willshire, A Data Quality Framework For Small Businesses, *Proceedings of the Eighth International Conference on Information Quality (ICIC-03).* 2003, pp: 239-244.

[4] Loshin, D. Data Quality ROI, *The Data Administration Newsletter*. 23.0, January 2003. http://www.tdan.com/i023fe02.htm.

[5] Orr, K. Data Quality and Systems, *Communications of the ACM*. 41(2) 1998, pp: 66-71.

[6] Pipino, L. L., Y. W. Lee, and R. Y Wang, Data Quality Assessment, *Communications of the ACM.* 45(4ve), pp: 211-218.

[7] Redman, T. C. The Impact of Poor Data Quality on the Typical Enterprise, *Communications of the ACM.* 41(2), 1998, pp: 79-82.

[8] Wang, R. Y. and D. M. Strong, Beyond Accuracy: What Data Quality Means to Data Consumers, *Journal of Management Information System.* 12(4) 1996, pp: 5-.33.

[9] Wang, R. Y, A Product Perspective on Total Data Quality Management, *Communications of the ACM.* 41(2) 1998, pp 58-65.

[10] Willshire, M. J. and D. Meyen, A Process for Improving Data Quality, *Data Quality*. 3(1) 1997. http://www.dataquality.com/997meyen.htm.